

## **Do firms compress the wage distribution?**

**Ana Rute Cardoso**

Institute for Economic Analysis (IAE-CSIC) and IZA Bonn

Campus UAB

08193 Bellaterra

Barcelona

Spain

[anarute.cardoso@iae.csic.es](mailto:anarute.cardoso@iae.csic.es)

July 2009

## **1. Introduction**

“Too many theories, too few facts.” This title, chosen by Baker and Holmstrom (1995) to their overview of the analysis of labour markets internal to the firm, remains today particularly suited to describe the work so far undertaken on the topic of the current chapter: Do firms compress the wage distribution relative to the distribution of worker productivities? If so, what is the impact of such compression on worker and firm performance and on labour turnover? A profusion of theories has not been matched by empirical testing, constrained as the analysis has been for a long time by data limitations. “Unfortunately, an appropriate data base with both worker and firm characteristics does not yet exist. A more complete understanding of the relationship between [...] productivity, and wage dispersion may have to await better sources of data” (Levine, 1991: 251). In the meantime, a long way has been tracked in the production and analysis of linked employer-employee data.

Early work on employer wages policies relied on extensive fieldwork on American companies. Interestingly, it called attention to the remarkable diversity in wage rates across firms, even when located in narrowly defined local labour markets, and the uniformity of wages inside the firm. Examples of this literature include Lester (1952), Dunlop (1957), Reynolds (1951) or, using econometric analysis on micro data on workers, Rees and Shultz (1970); an overview is provided by Kerr (1994). Ever since, the theoretical reasoning on why firms might have an incentive to compress the wage distribution relative to the distribution of worker productivities has been refined and some empirical testing using linked employer-employee data (henceforth LEED) has been undertaken.

This chapter starts with an overview of theoretical reasoning on the compression of wages by the firm. It aims, on one hand, at preparing the stage for

the review of the empirical literature that follows and, on the other hand, at helping define the setting for future empirical work, by getting into some detail on similarities and contrasts among the alternative theories. It then reviews empirical work that has used LEED to test theoretical predictions on the existence of a firm wage-compression effect and its implications on worker and firm productivity and labour turnover.

The chapter focuses on intra-firm wage dispersion, leaving out issues such as decomposition of inequality into its within- and between-firm components, differences between high- and low-wage firms or overall inequality in the labour market, to the extent that they are not directly linked to the level of wage inequality within the firm. The terms dispersion and inequality will be used interchangeably.

## **2. Theoretical framework on wage compression within the firm: a profusion of models**

A clarification of what is meant by compression of the wage distribution by the firm is in order. In a perfectly competitive labour market, workers' pay would be attached to their productivity and a single wage would hold for workers of the same quality, no matter which firm they would work for; on the contrary, workers of different quality would receive different wages, even if working for the same firm. However, the distribution of wages paid by each firm does not necessarily mirror the distribution of workers' productivities. To the extent that the distribution of workers' productivities within the firm is more stretched than that of workers' wages, we refer to the existence of a firm wage compression effect.

The key factors explaining the existence of an employer wage compression effect range from workers' preferences, to the production technology or the

information structure and frictions in the labour market, depending on the theoretical line considered.

Efficiency wage models in general are set in a framework of market frictions, such as hiring, training, firing and moving costs. Wages above the market clearing level would operate as a device to stimulate effort and prevent workers from shirking, discourage quitting that is costly to the firm given the existence of hiring, training and firing costs, or attract the more able workers, when they are heterogeneous in their ability and the wage the worker is willing to accept signals his/her ability. A positive link between the wage and worker effort therefore follows. The fairness or morale version of efficiency wages further brings into the analysis the level of dispersion of wages within the firm or the equity of its wage distribution. The assumption in this case is that there is interdependence of preferences and that the behaviour of the worker, in particular his effort, will depend on the norms of the group where he belongs and will drop if the wage level is perceived to be unfair. That would arguably be the reason why firms would adopt a consistent pay standard across their labour force – that is what is considered fair, and thus induces effort, when the utility of the worker is shaped also by equity and fairness considerations. It therefore follows that more unequal firm wage distributions would threaten the morale or the cohesiveness of the group and reduce productivity. Hamermesh (1975) introduced for the first time interdependencies among workers when modelling labour demand and labour supply, as individual effort and productivity would depend, not just on the worker's own wage, but also on a comparison wage, for example the average wage in the firm. Solow (1979) referred to adverse selection, shirking and morale considerations as keys to establish a link between wages and worker productivity, which would contribute to explain the slow adjustment of

wages when the macroeconomic conditions change. Akerloff (1982) and Akerloff and Yellen (1990) are widely quoted as the initial formulation of the morale efficiency wage theory in the economics literature. Levine (1991) stressed the mechanisms linking wage compression to worker cohesiveness (workers sharing and behaving by group norms) and its impact on firm performance. Alexopoulos and Cohen (2004) contribute to the morale efficiency wage literature by discussing the choice of comparison groups and modelling the inclusion of workers below the worker's own wage level in his/her comparison group, to conclude that firm's output and profits will decline if wage compression from below is imposed, since the level of effort of the better paid workers will decline. Skott (2005) highlights that norms change endogenously, as past events shape what is considered a fair—in other words, a normal— wage, which therefore adjusts to fit actual outcomes<sup>1</sup>, leading in turn to a sluggish adjustment of wages to labour market shocks.<sup>2</sup>

A different setting of interdependencies among preferences is assumed by Frank's (1984, 1985) theory on the quest for status—individuals care for their position in a hierarchy, comparing their own standing to the standing of others. Given distaste for being ranked low in a comparison group and given the assumption that individuals are free to choose the groups they join, in particular the firm they work for (the neighbourhood they live in, etc), they would presumably choose to associate with those who are similar to them and would sort into homogeneous groups. This is not the standard situation in the labour market, because some individuals are willing to pay for status and heterogeneity in willingness to pay

---

<sup>1</sup> Though it may also suffer exogenous shocks.

<sup>2</sup> This model is more concerned with wage changes than with wage levels.

renders status within one's hierarchy a tradable good like any other. Heterogeneous associations will thus form, as status seekers transfer resources to those who care less about status. The relevant comparison group would be the firm, a local hierarchy of co-workers, and the wage of workers of equal productivity would vary depending on the rank position: those occupying a high rank in a firm would have given up part of their wage (relative to their marginal productivity) to pay for status, while those occupying a low rank in a firm would be compensated with a wage premium (relative to their marginal productivity). This compensating differential would prevent workers at the bottom from moving to join a new firm, where the productivities and wages of co-workers would be closer to their own.<sup>3</sup> This trade-off between wage and status would be the key to explain the observed compression of the distribution of wages within firms, when compared to the distribution of worker productivities (Frank, 1984: 552).

An apparent contrast with the tournament theory is worth noting. Whereas according to Frank rank position within one's hierarchy matters and workers are willing to pay for it, some versions of tournament theory establish, on the contrary, a convex relationship between pay and hierarchical level —as one moves up to the top of the firm hierarchy, pay differences increase. In Lazear and Rosen (1981), such convexity of the pay scale would work as an incentive mechanism to illicit investment in skills and effort from the workers. Rosen (1986) introduces the option to move further up as an additional benefit to those who are promoted (beyond the immediate monetary gain) and as such, at the top of the hierarchy, an extra wage

---

<sup>3</sup> Frank assumes a competitive labour market and argues that on average it would still hold that workers in each firm would be paid their marginal productivity.

differential would be required to compensate for the lack of option to move further up. Nevertheless, Lazear (1989) considers the disruptive effects that tournaments may have on firm performance if workers can engage in counterproductive activities (such as sabotage of colleagues' work). In this case, he predicts some pay equality in profit-maximizing firms, to avoid disruptive behaviour by "hawks" in a system of pay for performance. One should however note that the type of pay equality Lazear refers to does not necessarily match that predicted by Frank. Frank's model considers the dispersion of the wage distribution relative to the dispersion of productivities, whereas in Lazear's model differences in productivity matter to the extent that the winner (the highest productivity worker) gets the pre-defined higher wage; as such, what matters in terms of productivity is just workers' ranking, and wage differences could be larger or smaller than realized productivity differences (see Lazear, 1984).

A model with interdependence of preferences and labour market frictions was proposed by Cabrales et al (2008), extending the dynamic setting of contract theory where, in the presence of uncertainty, risk-neutral firms offer contracts to their risk-averse workers. The model assumes that workers are heterogeneous, but perfect substitutes, and moreover that their true ability is unobserved, being revealed after the start of an employment spell. A particular type of inequity-aversion is considered —workers dislike inequality when they are the low-wage earners, but have no concern when they are the high earners. Their reference group includes colleagues in the same firm with similar skill (more precisely, those with “a similar career history within the firm”). As such, the comparison group changes endogenously, as workers enter and leave the group, due to job turnover. Interesting predictions emerge, concerning the link between productivity and wages and worker

segregation across firms. First of all, frictions in the labour market tend to reduce worker segregation across firms. Indeed, in the presence of moving or hiring costs, flows of workers across firms will be lower and, once the productivity of a worker inside a firm is revealed to be high, his/her wage does not increase as much —labour market frictions shield the employer from the competitive or outside wage rate—, but nevertheless s/he may stay with the firm. Within the firm, low and high skill workers will therefore co-exist. Secondly, workers' preference for equity will operate in the opposite direction, towards increasing segregation, simply because lower ability workers will suffer disutility from working with more skilled and higher paid colleagues and would prefer to join a firm with a more homogenous workforce. The relative importance of these two factors and the actual degree of segregation will depend on the extent of the moving costs. Thirdly, either factor operates to compress the wage distribution inside the firm relative to workers' productivities: labour market frictions enable the firm to appropriate part of the productivity gains from higher skilled workers, even though it is subject to outside competitive pressure when setting its wages; lower skilled workers will require a compensation for the disutility of working with higher paid colleagues and will thus receive productivity-unrelated wage increases.<sup>4</sup> Fourthly, the firm will use its personnel policy to manipulate the comparison group (for example, promoting some workers out of the comparison group and gradually increasing their salary, to reduce the costs of envy) and minimize the productivity-unrelated wage increases. The

---

<sup>4</sup> Moreover, in this setting of contract theory, insurance provision may further increase wage compression inside the firm.



dynamic nature of the model is a major accomplishment with respect to previous models.

Fehr and Schmidt (2001) and Sobel (2005) provide a very complete overview of models that assume the existence of interdependent preferences, distinguishing between stable preferences, when individuals care about the distribution of income or other payoffs at stake —they have “social preferences”, such as inequity aversion or altruism—, and context-specific preferences, when the individuals also care about the intentions behind the actions, the process or the environment —such as reciprocity models. Tests aimed at discriminating among these models have so far almost invariably relied on laboratory experiments and, less often, on field experiments.

Away from interdependent preferences, the model by Kremer (1993) assumes that technological reasons could induce each employer to systematically choose employees with certain homogeneous characteristics. In this case, interdependence of workers' skills is the crucial assumption. The skill of the worker is defined as the probability of successfully fulfilling his task and the basic idea is that the production process is made up of a series of complementary tasks, each of which can be performed perfectly or with errors. Only if all the tasks are completed perfectly will the product keep all its value and, in case of mistakes, the output will proportionately lose value —an analogy with the space shuttle *Challenger* is presented, where a small component among thousands of other ones, the O-rings, ruined the whole project. The output of a worker therefore depends, not just on his own skill, but as well on that of his co-workers. As such, the firm will have an incentive to combine workers of similar skill into the productive process: "firms with high  $q$  workers in the first  $(n-1)$  tasks place the highest value on having high-

skill workers in the  $n$ th task, so they bid the most for these workers. Thus, in equilibrium, workers of the same skill are matched together in firms." (Kremer, 1993: 554) Each firm is predicted to build a homogenous labour force, producing goods with corresponding quality, instead of hiring workers of different qualities and paying them the marginal product. The sorting of workers into firms according to their observed and unobserved ability would explain the wage differences among firms. Wages for different occupations would be correlated within firms due to the interactions among workers' skill or "multiplicative quality effects". Further developing the model generates the prediction that small differences in skill lead to sharp differences in wages and productivity across firms. Note however that in this case firms adopt a compressed wage distribution that reflects workers' ability (observed and unobserved) and their productivity, augmented by the presence of skill externalities.

Still other models allow for market imperfections as the key factor explaining the compression of wages but, different from models previously described, they make no assumption concerning interdependence of preferences. Acemoglu and Pischke (1999) explicitly consider labour market frictions that compress the distribution of wages relative to the distribution of productivities: collective bargaining wage floors, national minimum wages, asymmetric information between current and prospective employer, or job search and other mobility costs. Under any of these mechanisms, the outside wage option of an employed worker is lower than his/her productivity. For example, the worker may have to incur a mobility cost, such that, even if s/he were paid his/her marginal product in the new firm, the net benefit from changing firms would be lower; the possibility of an intervening period of unemployment also lowers the net benefit

from job changing; prospective employers cannot fully observe the worker's skill and therefore the wage offer may not fully reflect it. Any of these frictions gives the current employer some monopsony power, i.e. the ability to pay the worker a wage below his/her marginal productivity, as it takes into account the worker outside option. Moreover, most of these frictions mean that the rents the firm extracts from skilled workers are larger than those it extracts from the unskilled —if the cost of unemployment is larger for the skilled, in particular when the unemployment benefit system is progressive; when higher skills are harder to observe by prospective employers; or when minimum wages are enforced. Having pinpointed the reasons leading to wage compression inside the firm, these authors core analysis is devoted to showing why firms may have an interest in investing in worker training —as a result of training, the worker's productivity will increase but, given the compression of the wage distribution, the firm does not have to fully pay the worker for the productivity improvement.<sup>5</sup>

Manning (2003) provides a thorough discussion of monopsony models and several of its implications. In particular: the firm wage setting policy may detach wages from worker productivity, for example if the firm reacts to outside wage offers the worker may receive; worker turnover may be reduced if the firm matches outside wage offers, therefore resulting in lower turnover at the expense of higher

---

<sup>5</sup> Booth and Zoega (2004) further highlight that the firm may have an incentive to provide training even in situations usually associated with competitive wages and no market distortions, like piece-rate payments, as long as the worker does not receive fully the benefits of the increase in productivity.

internal wage dispersion; firms may be willing to provide and pay for training, in a type of reasoning with similarities to Acemoglu and Pischke (1999).

A few other bodies of theory were considered beyond the scope of this chapter. Insurance models per se, without any assumption on interdependent preferences, relate more directly to the reaction of wages to shocks and can explain why wages do not adjust as much as predicted by spot market theory. It is therefore more appropriate to analyze wage changes than wage levels. Incentives and delayed payment contracts, on the other hand, would lead one to expect the dispersion of wages to be larger than the dispersion of productivities, to the extent that young workers are paid below their marginal productivity and older workers are paid above their marginal productivity. Rent-sharing could explain wage differences across firms, as workers benefit from the good performance of their firm, but it does not provide an explanation why wage differentials inside the firm would be muted.<sup>6</sup>

### **3. Do firms compress the wage distribution?**

The challenge involved in testing the type of theories just described has been eloquently synthesized by Raff and Summers, even though they were focusing on efficiency wage theories alone: “The very impediments to evaluating workers’ ability, motivation and stability that might lead employers to pay efficiency wages make conventional testing of efficiency wage theories difficult. If the information needed to test these theories were available, there might be no need to pay efficiency wages. Econometric tests of efficiency wage theories also face the problem that variations in wages across firms or workers are unlikely to be exogenous,

---

<sup>6</sup> Rent-sharing is subject to detailed analysis in other chapters of this book.

complicating considerably the problem of identification." (Raff and Summers, 1987: S59) As such, tests on theories that assume interdependence of preferences have in the recent past often been conducted in laboratory experiments, which are not the scope of the current book. Instead, this chapter overviews empirical work using LEED that has aimed at testing theories that predict firm wage compression and, in some cases, attempted to disentangle explanations based on preferences, technology or market frictions.<sup>7</sup>

Machin and Manning (2004) concentrated on a narrowly defined occupation in a narrow geographic labour market (care assistants in elderly residential homes in a region in the UK), arguing that in this market one should find competitive wages, given the large number of firms delivering a homogenous good and the lack of union influence or minimum wage enforcement.<sup>8</sup> However, analysis of variance of wages indicates that wage dispersion across firms is large, while wage dispersion within the firm is small. That contrasts with the results on worker observable attributes (age, tenure, and hours worked), for which the within firm component of dispersion is very relevant. Corroborating evidence follows from the comparison of the determinants of wages and of the price of the product, taken as an indicator of performance, in two different samples: those firms that pay a flat wage to all of its workers, and those firms whose wage distribution has some degree of dispersion. Machin and Manning find that, while the determinants of the product price are similar across the two groups of firms, the determinants of wages are not,

---

<sup>7</sup> Away from the focus of this chapter on LEED, several studies provide a lively overview of the role of fairness in wage determination, based on real life situations and interviews with firm leaders.

<sup>8</sup> At the time the analysis was undertaken, there was no minimum wage in the UK.

concluding that worker attributes that are related to productivity are not the ones shaping wage differentials and, moreover, that worker unobservable quality cannot account for the differences in wages. Overall, they interpret their results as evidence against the competitive wage setting model, possibly linked to the existence of labour market frictions that grant employers some freedom when setting wages, combined with workers' preference for equity.

A narrow set of workers and firms has also been considered in Bishop's (1987) study. He relies on interviews with leaders of approximately 500 US firms, who provided information on two workers recently hired. The author runs regressions of the wage difference between these two workers on their productivity difference (as judged by supervisor rating) and a set of control variables. He concludes that relative wages inside the firm adjust to relative productivity, though by no means completely (as the elasticity is always below one). Larger firms adjust less their relative wages to productivity differences, which is interpreted as a result of higher monitoring costs, whereas firms in large markets adjust more, to avoid losing its more productive workers.

O'Reilly et al (1988) also relied on a narrow set of workers in approximately 100 US firms, using the wages of CEOs to explicitly test social comparison theories against tournament theory. They found that comparisons with a reference group outside the firm would lead, in this case, to increased inequality within the firm. In fact, they interpret the increasing wage differentials at the top of the hierarchy as evidence in favour of a social comparison model, after finding that the wages of CEOs were positively linked to the wages of outside directors in the board or the compensation committee. Members of the compensation committee would thus set wages having in mind social comparisons, with their own wages as a benchmark.

#### **4. Does wage compression lead to higher productivity and improved firm performance?**

Evidence on the impact of wage compression inside the firm on worker and firm performance is rather mixed, with results ranging from a positive impact to a negative one, and including a hump-shaped relationship or no significant relationship altogether.

Several studies on particular occupations and industries concluded that wage dispersion inside the firm leads to lower performance.

DeBrock et al (2004) used longitudinal linked employer-employee data to study the impact of wage dispersion on firm-level outcomes in a single industry — baseball teams. Though a particular sector, it enables a clearer observation of worker productivity and firm performance than most other sectors. Several measures of firm performance are considered, from won-lost records and attendance of games, to the value of the franchise, revenue and profits, but the two first ones are analysed, given the more reliable information available. The authors compute the expected or market wage for a particular worker, by regressing the salary on several worker attributes, from his age to statistics of his performance in the previous year and throughout his career. They then compute two measures of wage dispersion inside the firm using the Herfindahl index: the dispersion of expected wages, to capture the degree of heterogeneity in workers' observable quality; the dispersion in actual wages beyond the expected wage (residual wages), as an indicator of “unexplained” wage heterogeneity inside the firm. They find that wage dispersion has a negative impact on firm performance —teams with more homogenous worker observable quality

fare better and teams with less “unexplained” inequality also fare better<sup>9</sup>. Even though fairness considerations could play a role explaining the results, they interpret their evidence as more compellingly pointing to the relevance of matching workers with similar quality, given that the “technology” in this industry dictates a high degree of interdependence among workers’ skills. The matching of workers with similar quality would yield better outcomes, in a reasoning similar to that proposed by Kremer (1993).

Similarly, Bloom (1999) analysed the performance of baseball teams using longitudinal data both on players and teams. The dispersion measures are computed on raw wages and he considers different indicators of the team sportive and financial performance, as well as individual player performance. Results indicate that, after controlling for team effects and a wide set of worker attributes, higher dispersion of salaries within the team has a negative impact on team performance and on most measures of individual performance. Interdependencies among workers’ skill are again pointed out as the crucial factor leading to this outcome.

Another particular sector, that of academics, was analysed by Pfeffer and Langton (1993), relying on cross-section data on a wide sample of college and universities in the US. Despite the main focus on satisfaction, their work also deals with the impact of raw salary dispersion within a department on research productivity, measured as the number of publications authored<sup>10</sup>. They find that individual research productivity declines with salary dispersion within a given academic department.

---

<sup>9</sup> Though this last result is not always significant.

<sup>10</sup> Normalized by dividing by the average in the scientific field.



The quality of the product, as perceived by the customers, has been used by Cowherd and Levine (1992) as an indicator of performance for approximately 100 firms, mostly from the UK and US. They check the impact of inequality between the salary of top managers and lower level workers (hourly paid employees, as well as lower rank managers) on performance. The inequality measure basically compares the firm's wage distribution to the outside market wage distribution: first, the relative position that top managers in the firm occupy in the wage distribution of top managers in the same region or industry is computed; similarly for lower level workers, their pay ranking in the outside market is computed; subsequently, the ratio of the two provides an indication of the degree of inequity internal to the firm, when compared to the surrounding economy. After controlling for factors that may influence the pay of lower level workers or pay equity inside the firm, such as worker unionization, firm size and its market share, or the complexity of the production technology, the authors find that higher wage dispersion inside the firm leads to lower performance. The results are interpreted as evidence showing that pay equity boosts motivation and induces worker commitment to the firm's goals.

Studies relying on broader sets of occupations and industries, though controlling to different extents for these variables in their analyses, tend to find the opposite result, with a positive impact of wage dispersion on firm performance.

Swedish white-collar workers were the target of analysis of Heyman (2005), using longitudinal linked employer-employee data. Both the raw wage dispersion within the firm and unexplained wage dispersion are considered, the second one computed on the residuals of a wage regression run separately for each firm and year with the worker observable attributes (gender, education, experience, tenure) as explanatory variables. Similarly to DeBrock et al (2004), this measure captures

wage dispersion persisting after taking into account observable dimensions of the human capital of the workers. In a second stage regression, the firm performance is regressed on the first stage alternative measures of wage dispersion and control variables (such as composition of the workforce, industry, and firm size). Some versions of the analysis account for firm unobserved heterogeneity to address the problem of omitted variables (that may be correlated with wage dispersion) and instrument wage dispersion using its lagged values from four years earlier, to address the potential endogeneity problem. The author finds that wage dispersion among white-collars in the firm has a positive impact on firm performance.

Belfield and Marsden (2003) study performance-related pay and aim at judging on its relative merits, given that two opposite mechanisms may operate: an incentive mechanism, if it stimulates worker productivity; a disincentive mechanism, if equity is a matter of concern for workers. They use cross cross-section linked employer-employee data on the UK to check the impact of performance-related pay on both intra-firm wage inequality and on firm performance.<sup>11</sup> The financial performance of the firm is a self-reported measure, with establishment leaders asked to classify the establishment performance with respect to that of the other establishments in the same industry, from “much better” to “much worse” than average. The authors find that performance-related payment schemes are associated with both higher intra-firm inequality and better firm performance. They therefore conclude that any potential disincentive effects of performance-related pay are compensated by the incentive effects, resulting in a

---

<sup>11</sup> Note that workers can be rewarded according to their performance and still the distribution of wages can exhibit lower dispersion than the distribution of productivities, if the premium on productivity increases less than proportionally.

positive overall impact on firm performance. That is particularly true for firms who adopt the “right” kind of payment system (performance-related or input-related), as predicted by a wide set of variables characterizing their production regime, nature of jobs, degree of supervision, etc.

Lallemand, Plasman and Rycx (2004) also used cross-sectional linked employer-employee data, but on the Belgium economy. They run separate wage regressions for each firm and year and use the standard-error of each regression as the measure of wage dispersion inside the firm for similar workers. This is the explanatory variable of interest in a second-stage regression explaining firm performance (value-added per employee). Since wage dispersion may be endogenous, to the extent that good firm performance may lead to the award of wage premia to some workers, they instrument the dispersion of total salary with the dispersion of income taxes on salary excluding bonuses. Results indicate a positive impact of wage inequality inside the firm on firm productivity.<sup>12</sup>

Hibbs and Locking (2000) identify “good” and “bad” wage compression. Their reasoning departs from the changes that took place after mid 80s in the Swedish economy, as its wage setting system moved away from centralized collective bargaining with strong equity concerns and both the economy’s wage distribution and firms’ wage distributions grew more unequal. The authors estimate Cobb-Douglas production functions and labour productivity functions, either one augmented with the inclusion of two terms on wage dispersion: dispersion within-

---

<sup>12</sup> A positive impact of salary dispersion on firm performance tends to be found in other studies that concentrate only on the salary of managers and aim at testing tournament theories (see for instance Eriksson (1999) and Main *et al*, (1993)). Leonard’s (1990) work points to no significant relationship.

and between-firms. Using OLS to estimate the production functions and instrumental variables to estimate the productivity functions<sup>13</sup>, they reach two contrasting results: wage compression within the firm has an adverse effect on productive efficiency, reducing worker productivity and firm output, due to its (des)incentive effects, whereas wage compression across firms has a favourable impact on productive efficiency, as it provides an incentive for capital and labour to flow from less to more efficient firms. Though actually running their analysis at a rather aggregate level, these authors rely on variables computed from micro data on workers and firms.

An additional flavour to the results is brought by studies finding a hump-shaped relationship between firm wage dispersion and its performance.

Winter-Ebmer and Zweimuller (1999) use Austrian longitudinal linked employer-employee data and, lacking information on worker or firm performance, infer worker productivity from their wages. From a wage regression run separately for each firm and year, they retrieve: the estimated wage for a representative worker, which is taken as the indicator of firm productivity; the standard-error of the regression, taken as the measure of wage dispersion inside the firm for similar workers<sup>14</sup>. Subsequently, they run regressions of the proxy for productivity on a

---

<sup>13</sup> Output is an endogenous variable in the productivity regression, being instrumented with its own lagged values and the Swedish exports to OECD countries.

<sup>14</sup> They estimate tobit regressions, given the top-coding of wages, and include as independent variables age, gender, tenure, blue-collar, and foreigner status, but not education, which is not available.

quadratic term on wage dispersion and control variables<sup>15</sup>, finding a hump-shaped pattern in the relationship —some inequality inside the firm would improve productivity, but too much inequality would be detrimental. Nevertheless, assuming that firms pay their workers their marginal product (to justify the use of wages as a proxy for productivity) excludes the possibility that wages can be determined by equity or similar types of considerations. Under this setting, it is not clear why wage dispersion would have an impact on productivity and the capacity of the procedure to ascertain the relationship under testing is limited.

Linked employer-employee data on the Danish economy has been used by Bingley and Eriksson (2001), who similarly find, for white-collar workers, a hump-shaped relationship between wage dispersion and firm productivity —some inequality inside the firm improves performance, but too much inequality is counter-productive. They depart from a different methodology, though: firm performance is measured as total factor productivity (the Solow residual in a Cobb-Douglas production function)<sup>16</sup>; wage dispersion is computed as the standard deviation of residual wages (after controlling for worker observables); the equation of interest, estimated in the second stage, relates firm performance to wage dispersion. Also here instrumental variables are used, since good firm performance may be associated with the payment of wage premia to some types of workers (such as managers), thus increasing wage dispersion inside the firm. The study takes advantage of the

---

<sup>15</sup> Using OLS regressions on the contemporaneous levels of the dependent and independent variables and models with firm-specific effects, in a reasoning aimed at capturing longer-term relationships.

<sup>16</sup> They also proxy worker effort by the inverse of sickness absence (averaged for the firm level).

variation in the tax rates according to the worker municipality of residence, which impacts the dispersion of after-tax wages inside the firm but is assumed to have no direct impact on firm performance.

Finally, still another group of studies finds no significant relationship.

Arguing that wage increases, rather than wage levels, are judged by workers in terms of fairness and have a more direct impact on morale, Grund and Westergaard-Nielsen (2008) analyse Danish data on both wage levels and wage changes, limiting their analysis to workers who remain with the same firm for at least two periods. Their indicator of firm performance is value added per employee and wage dispersion is measured as the lagged coefficient of variation of wages. Relying initially on OLS regressions with several controls for the composition of the workforce, they find a hump-shaped relationship between the dispersion of wages inside the firm and its performance, consistent with the results by Winter-Ebmer and Zweimuller (1999). However that impact vanishes once unobservable differences across firms are accounted for by including firm fixed-effects in the regression.<sup>17</sup>

It is curious that studies considering a very narrow set of occupations tend to find a negative impact of inequality on performance, whereas studies relying on broader occupation groups tend to find a positive impact of inequality on performance. Under a setting of interdependent preferences, could it be that workers indeed adopt narrow comparison groups such as the specific occupation? Under a setting of interdependence of skills, could it be that studies using a broader set of

---

<sup>17</sup> When considering the dispersion of wage increases, instead, they find that for the bulk of firms and for white-collars in particular, fairness considerations dominate, with an increase in the dispersion of wage changes being associated with worse firm performance.

occupations are picking up positive externalities that may result from combining workers of different occupations within the firm?

Pfeffer and Langton's comment remains pertinent: "One of the more useful avenues for research on pay systems may be precisely this task of determining not which pay scheme is best but, rather, under what conditions salary dispersion has positive effects and under what conditions it has negative effects." (Pfeffer and Langton, 1993: 383)

### **5. Does wage compression reduce worker turnover?**

Evidence on the implications of firm wage compression on worker and job flows is rather scarce, but already divergent. In general, on theoretical grounds, morale and equity type of models would suggest that firms with higher inequality would have higher worker flows, whereas labour market frictions and monopsony type of models would lead to lower flows, also because employers would have more freedom to manipulate wages trying to retain certain workers.

The Slovenian economy is analysed by Haltiwanger and Vodopivec (2003) using longitudinal matched employer-employee data. They rely on several indicators of wage dispersion within the firm: i) raw wage dispersion; ii) residual wage dispersion, after controlling for worker observable attributes (education, experience and tenure); iii) similar residual wage dispersion, but further controlling for firm fixed effects; and finally iv) difference in wage dispersion between a model with firm and worker observable attributes plus their interaction, and a model with just worker observables. The aim is to pin down the dispersion resulting from idiosyncratic worker wage effects within the firm, given that: specification (ii) controls for the influence of collective bargaining that sets base wages relying on

worker observable attributes; specification (iii) further controls for firm-wide wage policies that affect every worker in the firm in the same way; specification (iv) retrieves wage dispersion within similar firms and within groups of workers with similar observable attributes, beyond the inequality existing in the overall economy within this group of workers. Subsequently, alternative measures of job and worker flows are regressed on these alternative indicators of wage dispersion (plus controls for the firm average residual wage and the industry). Results show that higher wage dispersion inside the firm<sup>18</sup> leads to less employment volatility, i.e. less job creation and destruction, which the authors interpret as an indication that, in the presence of more flexible wages, when shocks occur quantities do not need to adjust as much. However, results are less clear-cut once the analysis focuses on worker reallocation, since two opposing forces play a role. Note that excess worker reallocation, the major indicator commented upon, evaluates the flow of workers over and above what is strictly needed to account for the employment change in the firm, reflecting worker separations that are not due to job destruction and worker accessions that are not due to the need to fill a newly created job. Excess reallocation therefore quantifies matches worker-firm that are destroyed and replaced by another match, with no change in the overall employment level. Interestingly, the authors find that in firms with higher wage dispersion, excess worker separations come predominantly from worse matches (workers from the lower part of the residual wage distribution), and less so from better matches (workers from the higher part of the residual wage distribution). In other words, firms with high wage dispersion are able to retain their best workers, while inducing the worst ones to leave. As such,

---

<sup>18</sup> Lagged one period.



wage dispersion would work as a device used by firms to promote better job matches.

A comparable result is reached by Pfeffer and Davis-Blake (1992), though concentrating on a narrow occupational group, that of college and university administrators in the US. They find that a higher dispersion of the salary distribution is negatively associated with turnover for administrators with high salaries and positively associated with turnover for those with low salaries. Public knowledge of the salary distribution, either within the firm or in the external labour market, contributes to strengthen this relationship. In their analysis, the authors control for several other factors, both at the firm and the individual level, that may have an impact on turnover, such as the firm's size and financial resources, the type of institution and its funding source, the tenure structure of its workforce and the worker's gender and tenure.

A method to compute wage dispersion within the firm similar to Haltiwanger and Vodopivec (2003) had been used by Powell, Montgomery and Cosgrove (1994), as they relied on the residual wages retrieved from a regression of wages on worker observable attributes (in particular, education and experience) and establishment fixed-effects. Unlike Haltiwanger and Vodopivec, though, they concentrate on one single industry —child care centers in the US. They estimate tobit models on quit

and fire rates at the establishment level<sup>19</sup>, finding that wage dispersion within the firm has no significant impact on worker turnover.<sup>20</sup>

Mixed results are also reported by Heyman (2008), who uses Swedish longitudinal linked employer-employee data but nevertheless runs his analysis on industry level aggregates. Wage dispersion is computed as the coefficient of variation of raw wages within the industry. Alternative indicators of job flows —job reallocation and its separate components, job creation and job creation— are regressed on the wage dispersion measure. Industry unobservable heterogeneity is accounted for and the potential endogeneity of wage dispersion is handled by using its lagged values as instruments. Results show that in manufacturing, sub-industries with higher wage dispersion have lower job reallocation, in particular job destruction.<sup>21</sup> Again this author interprets the finding as suggesting that wage compression limits the extent of wage adjustments and therefore calls for employment adjustments once shocks hit an industry. However, in the services, the opposite effect is detected, as sub-industries with higher wage dispersion have

---

<sup>19</sup> Instrumenting the wage level of the establishment, which is also included in the regression, with indicators of the cost of living and the wages in the area, as well as the fees charged by the school.

<sup>20</sup> The analysis by Galizzi (2001) of the duration of employment spells, aimed at detecting contrasts in the behaviour of men and women, included among the explanatory variables the interaction of the worker's own wage with the average wage in the firm for the worker's broad occupation and gender. This measure of "relative wages" matches more closely an indicator of the firm wage level, rather than a measure of wage dispersion inside the firm.

<sup>21</sup> With no significant effect detected on job creation.

higher job reallocation<sup>22</sup>. Differences in the technology used and the possibly higher costs of job reallocation in the manufacturing sector may play a role explaining this difference in patterns between manufacturing and the services, but the exact mechanism in operation does not seem to have yet been pinpointed.

## **7. Conclusion**

Having in late-80s asked the question "Does the new generation of labor economists know more than the older generation?", Freeman asserted "the main conclusion I reach is that while labor economists are more knowledgeable of labor supply issues, we do not know more about firm behavior, labor demand and the overall functioning of the markets" (Freeman, 1989: 319). The empirical discussion on the existence of a firm wage compression effect and some of its implications is an example of the use of linked employer-employee data by the research community during the last couple of decades to gear research in the direction suggested by Freeman.

However, alongside with the potential of LEED, the difficulties involved in testing for the existence of a firm wage compression effect and its implications have been exposed. In fact, the margin of consensual results seems rather small: there is some degree of uniformity in wages within a firm, more so the narrower the set of occupations considered.

Once the aim is to explore the reasons and implications of an employer wage compression effect, we find all shades of results. Several reasons can justify this outcome: the data sets used have quite different characteristics, from coverage of a narrow set of occupations in a narrow set of firms, to coverage of the population of

---

<sup>22</sup> With no significant impact on job creation and job destruction separately.

firms and workers in an economy; the measures used, in particular on worker and firm performance, have also diverged widely; the empirical methods so far used are equally far apart in their capacity to tackle the issue and address the potential empirical problems. This strand of literature is probably still in its infancy (or teenagehood).

Identification of a test that would enable disentangling theories that often lead to similar predictions on this issue remains a challenge. One of the aims of this chapter has been to show some of the way that the empirical literature has covered and to highlight some of the major theories that could guide future empirical research on this topic.

## **References**

- Acemoglu, D. and J.S. Pischke (1999) "The structure of wages and investment in general training", *Journal of Political Economy*, 107(3), 539-572.
- Akerlof, George A. (1982) "Labor contracts as partial gift exchange", *Quarterly Journal of Economics*, 97(4), 543-567.
- Akerlof, George A.; Yellen, Janet L. (1990) "The fair wage-effort hypothesis and unemployment", *Quarterly Journal of Economics*, 105(2), 255-283.
- Alexopoulos, Michelle and Jon Cohen (2004) "What's wrong with forced wage compression? The fair wage hypothesis redux", *Economics Letters*, 83, 391-398.
- Baker, George and Bengt Holmstrom (1995) "Internal labor markets: Too many theories, too few facts", *American Economic Review*, 85(2), 255-259.

- Belfield, Richard and David Marsden (2003) "Performance pay, monitoring environments, and establishment performance", *International Journal of Manpower*, 24(4), 452-471.
- Bingley, Paul and Tor Eriksson (2001) "Pay spread and skewness, employee effort and firm productivity." Mimeo, University of Aarhus.
- Bishop, John (1987) "The recognition and reward of employee performance", *Journal of Labor Economics*, 59(4), S36-S56.
- Bloom, Matt (1999) "The performance effects of pay dispersion on individuals and organizations", *Academy of Management Journal*, 42(1), 25-40.
- Booth, A. L. and G. Zoega (2004) "Is wage compression a necessary condition for firm-financed general training?" *Oxford Economic Papers*, 56, 88-97.
- Cabrales, Antonio, Antoni Calvó-Armengol, and Nicola Pavoni (2008) "Social preferences, skill segregation, and wage dynamics", *Review of Economic Studies*, 75, 65-98.
- Cowherd, Douglas M. and David I. Levine (1992) "Product quality and pay equity between lower-level employees and top management: An investigation of distributive justice theory", *Administrative Science Quarterly*, 37, 302-320.
- Debrock, Lawrence, Wallace Hendricks, and Roger Koenker (2004) "Pay and performance: the impact of salary distribution on firm-level outcomes in baseball", *Journal of Sports Economics*, 5, 243-261.
- Dunlop, John T. (1957) "The task of contemporary wage theory", in his (ed.) *The Theory of Wage Determination*. London: MacMillan, 3-27.
- Eriksson, Tor (1999) "Executive compensation and tournament theory: empirical tests on Danish data", *Journal of Labor Economics*, 17(2), 262-280.

- Fehr, Ernst and Klaus Schmidt (2001) "Theories of fairness and reciprocity - evidence and economic implications", CEPR discussion paper 2703. London: Centre for Economic Policy Research.
- Frank, Robert H. (1984) "Are workers paid their marginal products?", *American Economic Review*, 74(4), 549-571.
- Frank, Robert H. (1985) *Choosing the Right Pond: Human Behaviour and the Quest for Status*. Oxford: Oxford University Press.
- Freeman, Richard B. (1989) "Does the new generation of labor economists know more than the older generation?" in his *Labour Markets in Action: Essays in Empirical Economics*. New York: Harvester Wheatsheaf, 317-342.
- Frick, Bernd, Joachim Prinz, and Karina Winkelmann (2003) "Pay inequalities and team performance: empirical evidence from the North American major leagues", *International Journal of Manpower*, 24(4), 472-488.
- Galizzi, Monica (2001) "Gender and labor attachment: Do within-firms's relative wages matter?", *Industrial Relations*, 40(4), 591-619.
- Grund, Christian and Niels Westergaard-Nielsen (2008) "The dispersion of employees' wage increases and firm performance", *Industrial and Labor Relations Review*, 61(4), 485-501.
- Haltiwanger, John and Milan Vodopivec (2003) "Worker flows, job flows and firm policies." *Economics of Transition*, 11(2), 253-290.
- Hamermesh, Daniel S. (1975) "Interdependence in the labor market", *Economica*, 42(168), 420-429.
- Heyman, Fredrik (2005) "Pay inequality and firm performance: evidence from matched employer-employee data", *Applied Economics*, 37, 1313-1327.

- Heyman, Fredrik (2008) "How wage compression affects job turnover", *Journal of Labor Research*, 29(1), 11-26.
- Hibbs, Douglas A. and Hakan Locking (2000) "Wage dispersion and productive efficiency: evidence for Sweden", *Journal of Labor Economics*, 18(4), 755-782.
- Kerr, Clark (1994) "The social economics revisionists: the 'real world' of labor markets and institutions." in Kerr, Clark; Staudohar, Paul D. (eds.) *Labor Economics and Industrial Relations: Markets and Institutions*. Cambridge, MA: Harvard University Press, 66-108.
- Kremer, Michael (1993) "The O-ring theory of economic development", *Quarterly Journal of Economics*, 108(3), 551-575.
- Lallemand, Thierry, Robert Plasman, and François Rycx (2004) "Intra-firm wage dispersion and firm performance: evidence from linked employer-employee data", *Kyklos*, 57(4), 451-566.
- Lazear, Edward P. (1984) "Incentives and wage rigidity", *American Economic Review*, 74(2), 339-344.
- Lazear, Edward P. (1989) "Pay equality and industrial politics", *Journal of Political Economy*, 97(3), 561-580.
- Lazear, Edward P. and Sherwin Rosen (1981) "Rank-order tournaments as optimal labor contracts", *Journal of Political Economy*, 89, 841-864.
- Leonard, Jonathan S. (1990) "Executive pay and firm performance", *Industrial and Labor Relations Review*, 43(3), 13S-29S.
- Lester, Richard A. (1952) "A range theory of wage differentials", *Industrial and Labor Relations Review*, 5, 483-500.

- Levine, David I. (1991) "Cohesiveness, productivity, and wage dispersion", *Journal of Economic Behavior and Organization*, 15, 237-255
- Machin, Stephen and Alan Manning (2004) "A test of competitive labor market theory: The wage structure among care assistants in the south of England", *Industrial and Labor Relations Review*, 57(3), 371-385.
- Main, Brian G. M., Charles A. O'Reilly III, and James Wade (1993) "Top executive pay: tournament or teamwork?" *Journal of Labor Economics*, 11(4), 606-628.
- Manning, Alan (2003) *Monopsony in Motion*. Princeton: Princeton University Press.
- O'Reilly III, Charles A., Brian G. Main, and Graef S. Crystal (1988) "CEO compensation as tournament and social comparison: A tale of two theories", *Administrative Science Quarterly*, 33, 257-274.
- Pfeffer, Jeffrey and Alison Davis-Blake (1992) "Salary dispersion, location in the salary distribution, and turnover among college administrators", *Industrial and Labor Relations Review*, 45(4), 753-763.
- Pfeffer, Jeffrey and Nancy Langton (1993) "The effect of wage dispersion on satisfaction, productivity, and working collaboratively: evidence from college and university faculty", *Administrative Science Quarterly*, 38(3), 382-407.
- Powell, Irene Mark Montgomery, and James Cosgrove (1994) "Compensation Structure and Establishment Quit and Fire Rates", *Industrial Relations*, 33(2), 229-248.
- Raff, Daniel M. G. and Lawrence H. Summers (1987) "Did Henry Ford pay efficiency wages?", *Journal of Labor Economics*, 5(4), S57-S86.



- Rees, Albert and George P. Shultz (1970) *Workers and Wages in an Urban Labor Market*. Chicago: University of Chicago Press.
- Reynolds, Lloyd G. (1951) *The Structure of Labor Markets*. Westport, CT: Greenwood Press, p. 155-229.
- Rosen, Sherwin (1986) "Prizes and incentives in elimination tournaments", *American Economic Review*, 76(4), 701-715.
- Sobel, Joel (2005) "Interdependent preferences and reciprocity", *Journal of Economic Literature*, 43(2), 392-436.
- Solow, Robert M. (1979) "Another possible source of wage stickiness", *Journal of Macroeconomics*, 1(1), 79-82.
- Skott, Peter (2005) "Fairness as a source of hysteresis in employment and relative wages", *Journal of Economic Behavior & Organization*, 57, 305–331.
- Slichter, Sumner H. (1950) "Notes on the structure of wages", *Review of Economics and Statistics*, 32(1), 80-91.
- Winter-Ebmer, Rudolf and Josef Zweimuller (1999) "Intra-firm wage dispersion and firm performance", *Kyklos*, 52, 555-572.